PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Sexual orientation identity in relation to smoking history and alcohol use at age 18/19: Cross-sectional associations from the Longitudinal Study of Young People in England (LSYPE)
AUTHORS	Hagger-Johnson, Gareth; Taibjee, Rafik; Semlyen, Joanna; Fitchie, Isla; Fish, Julie; Meads, Catherine; Varney, Justin

VERSION 1 - REVIEW

REVIEWER	Toby Lea, PhD Research Associate National Centre in HIV Social Research, The University of New South Wales, Australia
REVIEW RETURNED	19-Mar-2013

THE STUDY	Some further description of the sample would be good. At present there is gender, sexual orientation, ethnicity, and parental education and SES. Are there any of sample characteristics that could be described? There are some key references that should be included in the Introduction and Discussion sections. In the Introduction (para 2 and 3) the lit review around tobacco and alcohol is good, but lacking some key studies of tobacco and alcohol use among young people that have compared LGB and heterosexual young people. I would like to see more discussion of the findings of other studies of young people. This is particularly the case for alcohol. There are quite a few additional references that should be cited. In the Discussion section, page 13, line 14: reference 34 seems to be the wrong reference. Better references here would be work by Cochran et al; McCabe et al; Corliss et al. Page 14, para 1: there are some specific studies that have examined alcohol use in young people that should be cited here. This should emphasise the mixed findings that have been reported about alcohol use between LGB and heterosexual young people, particularly among young women. There is also no discussion of the findings of young men. The authors could consider including the following references, but a new search of the literature would also be a good idea to make sure key references have not been missed: e.g., Smith et al, 1999; Ziyadeh et al 2007; Russell et al, 2002; Eisenberg et al, 2003; Needham et al 2010.
GENERAL COMMENTS	This is an interesting paper, and one of few outside of the United States to report findings from a prospective cohort study of young people on sexual orientation and substance use.

	Some additional comments: Was illicit drug use measured? The authors should probably cite this as a limitation. Will data on drug use be collected in future survey rounds? What kind of data does the ONS Integrated Household Survey collect? Anything of relevance to the current paper? One issue with telephone and home interviews and LGB status is that parents being present may underestimate the % who were LGB because they had not yet disclosed to parents. I noticed that the authors made statistical comparisons about this, but could be stated more explicitly. In the Introduction, could you say "lowest quintile of schools" rather than "worst quintile....". Similarly, is there a synonym of "deprived" that could be used to describe schools? When discussing the findings on smoking in the Results and Discussion, the authors generally state that the data is smoking history, but sometimes this is not stated and could be confused for current smoking. Suggest always refer to smoking history or past smoking. Page 10, lines 35-41: there is a typo here. "And" is in the wrong place on line 36, should be on line 38. Page 10, line 50: the proportion (37.5%) is reported twice. I think this is a typo. I know that it isn't always necessary, but would the authors consider reporting p-values for odds ratios as well? Page 12, line 47: should be "and" between LGB and heterosexual. Page 13, line 5: does 1990/91 refer to year of birth? Should it not say 2004? Page 13, line 55: should say "A third limitation". Page 14, line 42: There is good evidence for minority stress, particularly the association between victimisation and substance use (see recent review by Collier et al 2013). Page 14, line 53-54: Im not sure what is meant by "a sexual preference involving tobacco". References: a consistent reference list style has not been used. Could you please check against author guidelines? In Table 1, there is a 'b' footnote next to ethnic minority status which looks like it should be removed. In Table 2, it should be stated somewhere in the table, caption or footer that the data represent ORs and 95% CIs.

REVIEWER	Matthias Wicki research fellow

	Research Institute of Addiction Switzerland Switzerland
REVIEW RETURNED	25-Mar-2013

THE STUDY	> Participants: the full sample at the beginning of the study was 15'770. the analytic sample was 6'656. why? (e.g. no reply at follow up?; missing values on a variable ? if so, on which one?) > Outcome variable: In the manuscript a person who drinks 7 times a week but gets drunk about every other time (i.e. is drunk 2-3 times a week) is coded as "non-hazardous alcohol use", while a person who drinks once a week and gets drunk every time (i.e. is drunk only 1 time a week) is coded as "hazardous alcohol use". this is very incomprehensible to me! additionally, the labels for the outcome variables are not consistent throughout the manuscript > Abstract etc: not complete and/or clear > Appropriate method: a linear regression to estimate the association between LG (0), bisexual (1) and heterosexual (2) identity and substance use (table 1) does not seem appropriate to me > English: English is not my first language (i hope my comments are helpful anyway). nonetheless, in a few places i asked myself, if the formulation would be understandable to others > References: i'm not up to date with the literature on the association between LGB identity and substance use. but e.g. an article by Bloomfield et al. which does report findings for the UK is missing Bloomfield et al (2011). International Differences in Alcohol Use According to Sexual Orientation, Substance Abuse, 32, 210–219
RESULTS & CONCLUSIONS	> Presentation: both tables could be presented in a more reader friendly way the result section in the manuscript (esp. for table 2) is very redundant with the table > Discussion: the discussion should refer more clearly to existing literature – also to explain the findings
REPORTING & ETHICS	> research ethics: based on the manuscript i do not see anything that is not appropriate. however, the article does not mention an ethical approval of the study. > publication ethics: based on the manuscript i do not see anything that is not appropriate. however, i'm not in position that i could affirm that there is no plagiarism or no undeclared conflict of interests.
GENERAL COMMENTS	The manuscript by Hager-Johnson and colleagues addresses an important topic: Is LGB identity associated with smoking and alcohol

	use? The study is especially important for two reasons: a) REPRESENTATIVE DATASET Unlike many of the early studies substance use among LGB (e.g. defined as identity, behaviour,) which were based on convenience samples (e.g. bars, prides) the study has a sample representative for the general population of the specific age group. b) DATA FROM OUTSIDE THE USA Most of the studies on the association between LGB and substance use are based on US-American samples; evidence from other countries are sparse. Even though substance use differs considerably between the USA and other countries (see e.g. findings from the GENACIS-project and the manuscript by Bloomfield et al, 2011) In my view the manuscript is based on a sound dataset, but could benefit considerably from some tidy-up-work. Besides three great concern there are quite some smaller flaws that might easily be revised. 1. ALCOHOL USE Unfortunately the LSYPE presumably did not include more common variables to measure alcohol use – esp. as far as "hazardous drinking" is concerned. Nonetheless, the way the two variables are treated does not allow to analyze their full information. a) weekly drinking is per se not a very good indicator risky alcohol use, neither as far as episodic [e.g. "binge"] nor chronic [e.g. WHO ">20/40g alcohol per day for men/women"]) is concerned. Why don't the authors use a more elegant method and try to predict the frequency of drinking? Mplus does offer nice models that consider more information than just transforming the frequency variable into a dichotomous « weekly vs non-weekly » variable! In case the authors want to stick to a very simple dichotomous variable, please consider to either give a good argument and reference for using the cut off. (In my view a cut off "drinking more than 3 times a week" i.e. not only drinking on weekends would be more informative). b) "Hazardous alcohol drinking" To begin with, the variable on "drinking to get drunk" might benefit from a clear label. I would suggest "risky single occasion drinking" (RSOD). "Intoxication" more likely to be used in a medical term (e.g. as a ICD10 diagnosis). Additionally in the manuscript meanders between "hazardous alcohol use", "hazardous drinking patterns", "hazardous alcohol drinking", this all makes it rather confusing. Often RSOD is defined as getting drunk / drinking more than a certain amount of alcohol twice a month or once a week. In the manuscript a person who drinks 7 times a week but gets drunk about every other time (i.e. is drunk 2-3 times a week) is coded as "non-RSOD", while a person who drinks once a week and gets drunk every time (i.e. is drunk only 1 time a week) is coded as "RSOD"! I'd assume that the authors did not do that intentionally! Or else, please give a rationale for doing so. The authors might want to combine the two variables in a way to

	measure RSOD as getting drunk once a week or more often: in a first step transfer the FREQUENCY variable into frequency per year (which I would recommend to do anyway, see comment “a”) and the information about getting drunk in a percentage (e.g. every time = 100%, most of the time = 75%) etc.). In a second step, combine frequency and prevalence and code cases over the cut off of 52 times per year as “RSOD”, these with less than 52 times as “non RSOD”. c) It is way more common to do analyses on alcohol use either among the full sample or among non-abstainers (i.e. having consumed alcohol at least once the past 12 months). I’d recommend doing the analysis on RSOD among non-abstainers. Thereby the results could easier be compared to other studies. If the authors do choose to do the analysis among weekly drinkers only, the rationale should be explicate in the method section or the implication for comparability should be mentioned in the discussion section. 2. REPORTING OF RESULTS Please report findings (especially the associations in table 2) separately for women and men! a) The preliminary analysis (test of an interaction effect) does not have much test power. b) Evidence from the literature shows that associations between LGB and substance use do vary by gender. (For a future meta-analysis, differences by gender would be way more valuable) c) There are not so many analyses in the manuscript (i.e. reporting the results by gender would increase the number of columns from 4 to six) 3. ETHNIC GROUP / SES Throughout the paper ethnic groups, SES etc is mentioned. However there is theoretical background to explain inasmuch this might be associated with the variables of interest. Also in the method section (e.g. sample weight, stepwise regression) and in the discussion these aspects are mostly missing. Please give a clear rationale. What is the reason to combine the three variables to “parental SES”? (p13 l19). Please elaborate in the introduction-section! In the stepwise regression models it is not very clear whether only for “parental educational attainment and occupational social class” (p12 l33) or for “education, occupation and ethnic minority status” (p22 l 24). Additionally: Are there any other variables in the dataset that might explain the association between LGB and substance use? e.g. “minority stress” is mentioned only in the discussion section, ... And now some minor comments following the order of the manuscript (page/line) 3/22 To me this sounds like during the home visits interviews and questionnaire have been used. But there were home interviews, telephone interviews and online questionnaires

	3/34 “Drinking alcohol to intoxication” please use here the same term as everywhere else in the paper 3/46 Please do not only report significant findings. The finding, that among women no association between a L or LB (vs. H) self identity and substance use was found, is important and noteworthy! “gay-identified”. Maybe it is due to my English but to me this sounds like “diagnosed” and not “men reporting a gay identiy”. 4/8 “Alcoholic intoxication” is rather a ICD10 diagnosis 5/15 The second article focus is incomprehensible to me – it “may be associated” does not say anything at all. Suggestion: previous results are mixed, therefore association should be tested. 5/28 WHERE ARE THE WOMEN!! Suggestion: “for men, but not for women ... “ 5/34 Well, the refusal rate of 0.1% is indeed very low. But according to social desirability it could be expected that there would be a pressure to say “hetero” as “i refuse to answer that question” already suggest that one might have a GLB identity 5/42 In my view the fact that it is a representative sample with a “credible” prevalence of LGB is a more important strength than that it is a longitudinal study (as no longitudinal analyses have been done) 5/45 Please be more precise: e.g. “smoking history (before the age of 16) was ... 7/17 Please be more specific than just say that estimates vary by age and ethnic group. Where is it high, where is it low? Does the estimate or the self-report vary? (Esp. the aspect of ethnic group would be interesting for the present paper) 7/29 “Cigarette smoking remains”: why “remain”? no trends have been reported in the manuscript. 7/45ff This paragraph is a bit messy. If there is a systematic review report, this should be reported at the beginning, and then later on going into details i.e. women, men. Please that studies that are already part of the review don’t “count twice”! Please do not only report the studies that found a significant association but report all studies! 7/47 “Modification by sex”

	I'd assume that the authors mean moderation instead of modification, and gender instead of sex 8/38 Please explain where about 6'000 adolescents that were not analysed in the study got lost! 8/45 Ethnic group, SES, "being social deprived" this is mentioned several times throughout the paper please elaborate why this is relevant. 8/53 Reporting the year and typical age is often a bit confusing. I'd suggest to introduce the terms t0, t1, t2 and indicate for each variable the when it was assessed 8/58 Please report here the percentage of home visit, telephone interviews and web questionnaires 9/1 Please indicate how LGB LG were coded 9/20 Please report the prevalence of LGB according to visit/phone/web. Only reporting that there was no significant difference (later on) is not enough (this is possibly due to a lack of test power). 9/24 "Regular cigarette smoking" – please use here the same label as used throughout the manuscript 9/45 Funny answer categories! Just a detail: are you sure that "almost every day" should not be "every day"? 10/12 if "parental SES" (consisting of three variables) should be used later on, please use here that term and explain it 10/15 Please explain how ethnic groups were coded (dummy variables?) 10/34 "Linear trends were" ... logistic regressions are not very useful to analyse linear trends! 10/49 The terms "minimally adjusted" and "fully adjusted" are not very common. I think you mean stepwise regression. e.g. "in the first step the models were adjusted for A B C and in the second step additionally for D E F". 11/48 Sorry I don't understand what the sentence "There were a lower proportion of lesbian women (37.5%) compared to gay men (37.5%), but a higher proportion of bisexual women (76.1%) than gay men (23.9%)." a) Lower proportion of what?

	b) What do the percentages in the brackets refer to and why is it the same for lesbian woman and gay men? c) Why comparing bisexual women to gay men d) “There was a lower proportion ... “ (not “there were”) 12/5 “A clear linear trend ... ” thank you for making me laugh! Please don’t consider a bisexual identity as the midpoint between LG and H on a interval scaled variable!! 11/23 “Complex survey design” is fix term used in the Mplus. If this is really what you mean, please introduce that in the method section 12/14 If the prevalence of self-reported LGB identity is different in according to visit/phone/web, why is this not considered (/adjusted for) in the analyses? 12/30ff The results reported here can be found exactly the same way in the table 2. This is not necessary. 12/24 The supplementary analyses should be done for all indicators of substance use and not only for smoking. If so: please provide the possibility that the reader may the analyses in table 2 separately by mode of administration. 13/43, 13/51 “50% more likely” (OR = 1.5) but “two-fold increase” (OR = 2.0)? this is a bit confusing (even though it might be correct English). 13/56 Again: Please don’t just focus on significant results! There has also no significant association for B or LB women be found. 14/2ff The paragraph on “strengths” may be shortened Instead please elaborate on the integration of the findings in the existing research literature. 15/35 “We are not aware”: please have a look at the literature and if nothing can be found, then please write so. The fact that you are “not aware” is less relevant here. Table 1 The table is quite confusing: e.g. where can are there N(%) and where mean (SD) [i think there are no means at all!?], the % are sometimes row-% sometimes column-% Suggestion: - If the n for LG B H are reported, it has not to be reported a second time in the column, percentages are enough - “N” is used for the total sample, “n” is used for sub-sample - Please report in a column when the variable was measured (e.g.

	t2) or at what age - Please report three decimals for the “p” - Please report differences for LG vs. H, B vs. H and LGB vs. H (this is the same structure as later on in table 2) - “n = 6656” is not the column header for the variables below In the remarks for table 2 it says that alcohol use was measured in 2004 and 2009. I’d assume that only the data from 2009 has been used! If not: please mentioned that in the method section Also in the remarks it says that smoking was measured in 2004! I hope this is just a typo. Table 2 - Title should mention smoking and alcohol - The headers of the column should mention the same variable label as throughout the paper - Instead of minimally/fully adjusted -> “stepwise”

VERSION 1 – AUTHOR RESPONSE

Reviewer 2

R2.1. The full sample at the beginning of the study was 15,770. The analytic sample was 6,656. Why? (e.g. no reply at follow up?; missing values on a variable ? if so, on which one?)

Response: The study design is a prospective cohort study with repeated follow-ups. It is common for the sample size to gradually decrease over time. Sexual orientation was only introduced at wave 6, and therefore the available data refer to wave 6 rather than at baseline (15,770). Full details are provided in the reference to the study documentation so that readers can find out more about the sample sizes at each wave.

R2.2. In the manuscript a person who drinks 7 times a week but gets drunk about every other time (i.e. is drunk 2-3 times a week) is coded as “non-hazardous alcohol use”, while a person who drinks once a week and gets drunk every time (i.e. is drunk only 1 time a week) is coded as “hazardous alcohol use”. This is very incomprehensible to me!

Response: Following the reviewer’s excellent suggestion to use ‘risky single occasion drinking’ as an outcome, instead of frequency of drunkenness, we have repeated the analyses.

R2.3. Additionally, the labels for the outcome variables are not consistent throughout the manuscript.

Response: We have now used consistent labels throughout the manuscript: smoking history, drinking alcohol >2 days per week, and risky single occasion drinking.

R2.4. Abstract not complete and/or clear.

Response: We have attempted to make the abstract clearer, but were not sure in what way it is incomplete. We welcome further suggestions on how to make it clearer and complete.

R2.5. A linear regression to estimate the association between LG (0), bisexual (1) and heterosexual (2) identity and substance use (table 1) does not seem appropriate to me.

Response: Linear or logistic regression can be used to test for trend, by including a semi-continuous predictor and a continuous/logistic outcome for the variable concerned. This produces the same p value as a test for linear trend. In any case, following the objections from reviewers on treating bisexual as a midpoint between LG and heterosexual, we have removed this test.

R2.6. English is not my first language (I hope my comments are helpful anyway). Nonetheless, in a few places I asked myself, if the formulation would be understandable to others.

Response: The reviewer's comments are extremely helpful and we have tried to improve the readability of the manuscript for the widest possible audience.

R2.7. I'm not up to date with the literature on the association between LGB identity and substance use. But e.g. an article by Bloomfield et al. which does report findings for the UK is missing Bloomfield et al (2011). International Differences in Alcohol Use According to Sexual Orientation, Substance Abuse, 32, 210–219.

Response: We have now included this study in the introduction section.

R2.8. Both tables could be presented in a more reader friendly way the result section in the manuscript (esp. for table 2) is very redundant with the table.

Response: We have modified the tables to make them more accessible, for example, by removing the p value for linear trend and presenting the results for the whole sample rather than for men and women separately (following the suggestion that the study is under-powered to evaluate effect modification).

R2.9. The discussion should refer more clearly to existing literature – also to explain the findings.

Response: We have modified the discussion to refer more clearly to existing literature, helping to explain the findings.

R2.10. based on the manuscript I do not see anything that is not appropriate. However, the article does not mention an ethical approval of the study.

Response: Our study involves analysis of a publicly available and anonymized dataset, and therefore does not require ethical approval. Relevant permissions to collect the data were obtained by the LSYPE team at prior to data collection.

R2.11. Based on the manuscript I do not see anything that is not appropriate. However, I'm not in position that I could affirm that there is no plagiarism or no undeclared conflict of interests.

Response: We are happy for the manuscript to be checked against any plagiarism detection software for evidence of plagiarism. All authors will submit declarations of potential conflicts of interest individually.

R2.12. The manuscript by Hagger-Johnson and colleagues addresses an important topic: Is LGB identity associated with smoking and alcohol use? The study is especially important for two reasons:
a) REPRESENTATIVE DATASET
Unlike many of the early studies substance use among LGB (e.g. defined as identity, behaviour,) which were based on convenience samples (e.g. bars, prides) the study has a sample representative for the general population of the specific age group.
b) DATA FROM OUTSIDE THE USA
Most of the studies on the association between LGB and substance use are based on US-American samples; evidence from other countries are sparse. Even though substance use differs considerably between the USA and other countries (see e.g. findings from the GENACIS-project and the manuscript by Bloomfield et al, 2011)

Response: We are extremely grateful for the reviewer's positive comments. We appreciate that ours results differ from the pooled international analysis, but emphasize the importance of publishing data from a young cohort of English people, not previously captured by prior studies. Cohort effects may have produced different associations in this young age group. We hope that our results can be pooled into future meta-analyses, adding to the evidence base.

R2.13. In my view the manuscript is based on a sound dataset, but could benefit considerably from

some tidy-up-work. Besides three great concern there are quite some smaller flaws that might easily be revised.

Response: We are thankful for the suggestions below, which have been incorporated into the revision.

R2.14. Unfortunately the LSYPE presumably did not include more common variables to measure alcohol use – esp. as far as "hazardous drinking" is concerned. Nonetheless, the way the two variables are treated does not allow to analyse their full information.

Response: We agree that we did not make full use of the two alcohol variables, and very much appreciate the suggestions below.

R2.15. a) weekly drinking is per se not a very good indicator risky alcohol use, neither as far as episodic [e.g. "binge"] nor chronic [e.g. WHO ">20/40g alcohol per day for men/women"]) is concerned. Why don't the authors use a more elegant method and try to predict the frequency of drinking? Mplus does offer nice models that consider more information than just transforming the frequency variable into a dichotomous « weekly vs non-weekly » variable! In case the authors want to stick to a very simple dichotomous variable, please consider to either give a good argument and reference for using the cut off. (In my view a cut off "drinking more than 3 times a week" i.e. not only drinking on weekends would be more informative).

Response: We acknowledge that frequency of alcohol drinking is not a good indicator of risky alcohol use per se (although more frequent drinkers tend to drink more heavily). In line with the suggestion however, we have rerun the analysis using drinking more than twice per week, which we think captures participants drinking on days additional to Friday/Saturday evenings – the typical pattern in this age group (Measham & Brain, 2005, Crime, Media and Culture).

R2.16. b) "Hazardous alcohol drinking". To begin with, the variable on "drinking to get drunk" might benefit from a clear label. I would suggest "risky single occasion drinking" (RSOD). "Intoxication" more likely to be used in a medical term (e.g. as a ICD10 diagnosis). Additionally in the manuscript meanders between "hazardous alcohol use", "hazardous drinking patterns", "hazardous alcohol drinking", this all makes it rather confusing. Often RSOD is defined as getting drunk / drinking more than a certain amount of alcohol twice a month or once a week.

Response: We have replaced this outcome with RSOD (see below) and used a consistent label throughout the manuscript.

R2.17. In the manuscript a person who drinks 7 times a week but gets drunk about every other time (i.e. is drunk 2-3 times a week) is coded as "non-RSOD", while a person who drinks once a week and gets drunk every time (i.e. is drunk only 1 time a week) is coded as "RSOD"! I'd assume that the authors did not do that intentionally! Or else, please give a rationale for doing so.

Response: We agree that our previous definition did not make full use of the data available. We have rerun the analysis using RSOD (see below) instead.

R2.18. The authors might want to combine the two variables in a way to measure RSOD as getting drunk once a week or more often: in a first step transfer the FREQUENCY variable into frequency per year (which I would recommend to do anyway, see comment "a") and the information about getting drunk in a percentage (e.g. every time = 100%, most of the time = 75%) etc.). In a second step, combine frequency and prevalence and code cases over the cut off of 52 times per year as "RSOD", these with less than 52 times as "non RSOD".

Response: We would like to thank the reviewer for this excellent suggestion which we have adopted.

R2.19. It is way more common to do analyses on alcohol use either among the full sample or among non-abstainers (i.e. having consumed alcohol at least once the past 12 months). I'd recommend doing the analysis on RSOD among non-abstainers. Thereby the results could easier be compared to other studies. If the authors do choose to do the analysis among weekly drinkers only, the rationale should be explicate in the method section or the implication for comparability should be mentioned in the

discussion section.

Response: We have now used a single and larger analytic sample for all analyses, which includes abstainers. In sensitivity analyses, we reran the models after excluding abstainers and report these results in the paper.

R2.20. Please report findings (especially the associations in table 2) separately for women and men!

Response: Following concerns raised by reviewers that power is low for testing for effect modification by gender, we have presented results on the combined sample. Separate results for men and women are supplied as a supplementary table, but with a cautionary note that chance findings may produce significant associations.

R2.21. The preliminary analysis (test of an interaction effect) does not have much test power.

Response: We agree that power may be too low to meaningfully test for effect modification by sex and socio-economic status. Therefore, we have presented results for the entire analytic sample (e.g. men and women, controlling for sex), reducing the likelihood of chance findings.

R2.22. Evidence from the literature shows that associations between LGB and substance use do vary by gender. (For a future meta-analysis, differences by gender would be way more valuable).

Response: We acknowledge that there are sex differences in the association between LGB identity and substance use, which we hope is clearer in the introduction. The pooled analysis of international differences in alcohol use findings helps to emphasize this point. Given that our study has lower power, our view is that the main results should combine males and females, but supplementary analyses should be used to illustrate possible sex differences. We hope that our results can be pooled into future meta-analytic work.

R2.23. There are not so many analyses in the manuscript (i.e. reporting the results by gender would increase the number of columns from 4 to six)

Response: We would like to keep the results as simple as possible. Therefore, the main table shows results for males and females; supplementary analyses separate the two groups.

R2.24. Throughout the paper ethnic groups, SES etc is mentioned. However there is theoretical background to explain in as much this might be associated with the variables of interest. Also in the method section (e.g. sample weight, stepwise regression) and in the discussion these aspects are mostly missing. Please give a clear rationale.

Response: We have added ‘Ethnicity and socio-economic status are possible confounding factors, because they may be associated both with sexual identity and with health behaviours’.The way in which ethnicity and SES are used in sampling weights is a separate issue. Ethnic minorities and schools with higher levels of socio-economic deprivation were over-sampled at recruitment. Allowing for sampling weights provides correct standard errors in the analysis. We hope that this is now more clear in the revised version of our paper.

R2.25. What is the reason to combine the three variables to “parental SES”? (p13 l19). Please elaborate in the introduction-section!

Response: Given space limitations, we have added a brief explanation to the methods section: ‘Parental education attainment and occupational class are both considered indicators of parental socio-economic status (SES)’

R2.26. In the stepwise regression models it is not very clear whether only for “parental educational attainment and occupational social class” (p12 l33) or for “education, occupation and ethnic minority status” (p22 l 24).

Response: We are grateful to the reviewer for pointing this out. In the revision, it is made clear that the minimally adjusted model refers to age and sex, the fully adjusted model refers to additional

adjustment for ethnic minority status and parental SES (education and occupation).

R2.27. Additionally: Are there any other variables in the dataset that might explain the association between LGB and substance use? e.g. “minority stress” is mentioned only in the discussion section, ...

Response: Our paper is intended to describe rather than explain the association, and is therefore a first step in an intended larger programme of research. We prefer to discuss possible mechanisms for future study, rather than attempt to explain the association within the same paper.

R2.28. And now some minor comments following the order of the manuscript (page/line)

Response: We are very grateful to the reviewer for making such detailed comments.

R2.29. (3/22) To me this sounds like during the home visits interviews and questionnaire have been used. But there were home interviews, telephone interviews and online questionnaires

Response: We have now referred to ‘Self-completion questionnaires during home visits, face-to-face interviews, and web-based questionnaires’.
R2.30. (3/34) “Drinking alcohol to intoxication” please use here the same term as everywhere else in the paper

Response: This is now replaced with risky single occasion drinking.

R2.31. (3/46) Please do not only report significant findings. The finding, that among women no association between a L or LB (vs. H) self identity and substance use was found, is important and noteworthy!
“gay-identified”. Maybe it is due to my English but to me this sounds like “diagnosed” and not “men reporting a gay identity”.

Response: By ‘gay-identified’ we meant ‘gay self-identified’ but have removed this in the new version.

R2.32. (4/8) “Alcoholic intoxication” is rather a ICD10 diagnosis

Response: See above, this term is no longer used.

R2.33. (5/15) The second article focus is incomprehensible to me – it “may be associated” does not say anything at all. Suggestion: previous results are mixed, therefore association should be tested.

Response: We have changed this to ‘Previous results for alcohol use are mixed, therefore the association should be tested’

R2.34. (5/28) Where are the women!! Suggestion: “for men, but not for women ... “

Response: This is not applicable in the revised paper.

R2.35. (5/34) Well, the refusal rate of 0.1% is indeed very low. But according to social desirability it could be expected that there would be a pressure to say “hetero” as “i refuse to answer that question” already suggest that one might have a GLB identity.

Response: We agree with the reviewer, and have made this point more salient. The heterosexual category is likely to contain LGB people who feel pressure to answer as heterosexual. This will introduce misclassification bias, so we have probably under-estimated the true association: ‘Although the refusal rate for the question was low, some participants who identify as LGB might have responded ‘heterosexual’ for this and other reasons, which include socially desirable responding. This would lead to misclassification bias, leading us to have under-estimated the size of any associations found’

R2.36. (5/42) In my view the fact that it is a representative sample with a “credible” prevalence of LGB is a more important strength than that it is a longitudinal study (as no longitudinal analyses have been

done).

Response: We agree and have emphasized that our study was essentially cross-sectional analysis from a longitudinal study. We want to emphasize the importance of future longitudinal follow-ups however, and have added: 'In our view, there is a clear need for a repeated measures cohort study of LGB people'

R2.37. (5/45). Please be more precise: e.g. "smoking history (before the age of 16) was ...

Response: We have added to the methods section 'Cigarette smoking was last measured in 2006 (typical age 15/16)'

R2.38. (7/17). Please be more specific than just say that estimates vary by age and ethnic group. Where is it high, where is it low? Does the estimate or the self-report vary? (Esp. the aspect of ethnic group would be interesting for the present paper)

Response: We have added 'Estimates tend to be larger when attraction or sexual behaviour are used to define sexual orientation' to the introduction, and discussed ethnic/SES differences (see point above).

R2.29. (7/29). "Cigarette smoking remains": why "remain"? No trends have been reported in the manuscript.

Response: we meant that over time, smoking has decreased among young people in the UK, but it still remains a problem. For clarity we have changed this to 'Cigarette smoking is a prevalent behaviour among young people'

R2.30.(7/45ff). This paragraph is a bit messy. If there is a systematic review report, this should be reported at the beginning, and then later on going into details i.e. women, men. Please that studies that are already part of the review don't "count twice"!
Please do not only report the studies that found a significant association but report all studies!

Response: We have modified the structure of this paragraph and removed the majority of studies that are already part of the review articles or the pooled analysis.

R2.31. (7/47). "Modification by sex". I'd assume that the authors mean moderation instead of modification, and gender instead of sex.

Response: 'Effect modification' (more commonly used in epidemiology) is the same as 'moderation' (more common in behavioural sciences). We have changed to 'gender' when described, to avoid confusion with sexual identity or behaviour.

R2.32. (8/38). Please explain where about 6000 adolescents that were not analysed in the study got lost!

Response: See above. Our study uses data from wave 6 which has a smaller sample size than at recruitment. It is common for sample sizes to reduce over time in repeated measures longitudinal research.
R2.33. (8/45). Ethnic group, SES, "being social deprived" this is mentioned several times throughout the paper please elaborate why this is relevant.

Response: See above – ethnicity/SES are possible confounding factors, schools with socio-economic deprivation were over-sampled.
R2.34. (8/53). Reporting the year and typical age is often a bit confusing. I'd suggest to introduce the terms t0, t1, t2 and indicate for each variable the when it was assessed.

Response: In order to simplify presentation and allow comparability across studies and help future pooled analyses, we are keen to use years and ages. These have greater meaning to the widest possible range of readers. Using t0, t1, t3 refers only to the current study design and would lack meaning outside of this study; many readers would find it too technical.

R2.35. (8/58). Please report here the percentage of home visit, telephone interviews and web questionnaires

Response: We have added this information to the statistical analysis section of the manuscript: interview (12.2%), telephone interview (47.6%), web questionnaire (40.1%)

R2.36. (9/1). Please indicate how LGB LG were coded.

Response: This depended on the analysis. For Model 1, LG were coded 1 vs. 0, B were coded 1 vs. 0 (reference category = heterosexual). For Model 2, LGB were coded 1 vs. 0.

R2.37. (9/20). Please report the prevalence of LGB according to visit/phone/web. Only reporting that there was no significant difference (later on) is not enough (this is possibly due to a lack of test power).

Response: We agree that this test is under-powered, and have added 'there were only 33 LGB participants for interviews, 86 for telephone interviews and 115 for web questionnaires' to make this point clear.

R2.38. (9/24). "Regular cigarette smoking" – please use here the same label as used throughout the manuscript.

Response: We have used 'smoking history' throughout the manuscript.

R2.39. (9/45). Funny answer categories! Just a detail: are you sure that "almost every day" should not be "every day"?

Response: We agree with the reviewer, but were not involved ourselves in the development of these questions.

R2.40. (10/12). If "parental SES" (consisting of three variables) should be used later on, please use here that term and explain it.

Response: At first usage, we explain the parental SES refers to occupational social class and education.

R2.41. (10/15). Please explain how ethnic groups were coded (dummy variables?)

Response: All ethnic minorities were grouped into ethnic minority status (1) versus white (0): 'White, Mixed, Indian/Pakistani/Bangladeshi, Black Caribbean/Black African, Chinese; grouped into 'ethnic minority' (1) or White (0)'

R2.42. (10/34). "Linear trends were" ... logistic regressions are not very useful to analyse linear trends!

Response: Using a semicontinuous variable (0/1/2) as a predictor variable where the variable of interest is the dependent variable, produces the same p value as a standard test of linear trend. Following the request not to treat LG/B/hetero as a linear sequence, this has been removed.

R2.43. (10/49). The terms "minimally adjusted" and "fully adjusted" are not very common. I think you mean stepwise regression. e.g. "in the first step the models were adjusted for A B C and in the second step additionally for D E F".

Response: Minimally adjusted is a term widely used in epidemiology, usually referring to age/sex. We have made it clearer which variables were used in each model, in the revision.

R2.44. (11/48). Sorry I don't understand what the sentence "There were a lower proportion of lesbian women (37.5%) compared to gay men (37.5%), but a higher proportion of bisexual women (76.1%) than gay men (23.9%)."

Response: For clarity, we have changed this to ‘There were fewer women (37.5%) in the LG category than men, but more women (76.0%) in the B category than men’

R2.45. (12/5). “A clear linear trend ... ” thank you for making me laugh! Please don’t consider a bisexual identity as the midpoint between LG and H on a interval scaled variable!!

Response: This has now been removed as requested.

R2.46. (11/23). “Complex survey design” is fix term used in the Mplus. If this is really what you mean, please introduce that in the method section.

Response: We have changed this to ‘When using sample weights to correct for over-sampling of ethnic minority groups and schools with higher socio-economic deprivation however, the proportion of participants classified as LGB was 3.5% (1.3 LG, 2.2% B)’.

R2.47. (12/14). If the prevalence of self-reported LGB identity is different in according to visit/phone/web, why is this not considered (/adjusted for) in the analyses?

Response: There are small numbers of home visit interviews. Additionally, the mode of survey administration is unlikely to be associated with smoking or alcohol use. Therefore, it should not produce a biased estimate if we do not control for survey mode. We did however, run additional analysis controlling for survey mode and results did not differ materially (not shown, available on request). This mitigates concerns that the results are driven by survey mode artefacts.

R2.48. (12/30ff). The results reported here can be found exactly the same way in the table 2. This is not necessary.

Response: We have removed duplication of results in the results and tables.

R2.49. (12/24). The supplementary analyses should be done for all indicators of substance use and not only for smoking. If so: please provide the possibility that the reader may the analyses in table 2 separately by mode of administration.

Response: We have removed these analyses following concerns that they are underpowered, and made a note in the discussion that survey mode is not likely to be associated with the outcomes of interest. We undertook additional analyses after controlling for survey mode (not shown, available on request) and the results were not changed materially.

R2.50. (13/43, 13/51). “50% more likely” (OR = 1.5) but “two-fold increase” (OR = 2.0)? this is a bit confusing (even though it might be correct English).

Response: We have made the language consistent throughout the manuscript.

R2.51. (13/56). Again: Please don’t just focus on significant results! There has also no significant association for B or LB women be found.

Response: We have now described all the results, including non-significant ones.

R2.52. (14/2ff). The paragraph on “strengths” may be shortened. Instead please elaborate on the integration of the findings in the existing research literature.

Response: We have elaborated further on how the findings contribute to the existing literature, but think that the points made in the strengths paragraph are important enough to remain. It is important to emphasize the representative nature of the sample, the low refusal rate, and the fact that the UK is less placed to provide data. Previous studies have used surveys/convenience samples such as pride events and so our study is a major and rare contribution.

R2.53. (15/35). “We are not aware”: please have a look at the literature and if nothing can be found, then please write so. The fact that you are “not aware” is less relevant here.

Response: We have changed to ‘apparently no evidence’. It is important to acknowledge that there may be evidence not captured in online databases (e.g. grey literature, conference talks, unpublished work).

R2.54. The table is quite confusing: e.g. where can are there N(%) and where mean (SD) [i think there are no means at all!?], the % are sometimes row-% sometimes column-%

Response: We have improved the table to make this clear.

R2.55. - If the n for LG B H are reported, it has not to be reported a second time in the column, percentages are enough

Response: We are not sure what the reviewer means, but hope that the table is now clearer.

R2.56. Other comments about Table 1.

Response: We have used N/n for total/sub sample. We have added the age at which variables were measured (if not specified, these were at recruitment). We have used three decimals for the p value. The p values now follow the same structure as the analyses in Table 2. The column header is now ‘study variables’ instead of the N.

R2.57. Table 2.

Response: The title is now ‘Association between sexual orientation identity and smoking, alcohol drinking >2 days/week and risky single occasion drinking’. As noted above, we prefer to keep minimally/fully adjusted but have made it clearer what this means.

VERSION 2 – REVIEW

REVIEWER	Dr Toby Lea Research Associate Centre for Social Research in Health, the University of New South Wales, Australia
REVIEW RETURNED	02-Jul-2013

THE STUDY	While the literature review is greatly improved, in the Introduction section some of the references for alcohol use reference papers that appear to be about tobacco (eg. references 3 and 20). Could the authors please check all references in the body of the manuscript to make sure they correspond with the correct reference in the reference list. Sometimes if multiple authors are accessing the same document, Endnote can do strange things and change the references. This is what may have happened here.
GENERAL COMMENTS	The authors have done a great job at revising this manuscript. A couple of small things: 1. In the Abstract one of the odds ratios appears to report the lower bound of the 95% CI as the odds ratio (bisexual identity associated with RSOD (OR 1.02 95% CI 1.02). Please check this one. 2. Article focus: regarding the second point about alcohol, please spell out what is mixed. While I get what the authors mean, it is unclear. 3. In the Methods section, I don't understand what is meant by "ethnic minority groups to N=1000 per group". Could the authors please explain this a bit more clearly if possible? 4. In the Results, there is a % sign missing after 1.3 LG, 2.2% GB. 5. I would avoid using 'B' and 'L' and 'G' on their own and would opt

	instead to spell these words out when used on their own. LG and GB and LGB are fine however. 6. In the Results, there are several instances where the authors should spell out who they are comparing LG and bisexual respondents to so it is more clear. e.g. "Lesbian and gay participants were more than twice as likely to have a history of cigarette smoking compared to heterosexual participants". 7. Were any comparisons made between LG and bisexual participants? If not, why not? Could the authors please cite as a limitation. I suspect it is because of a lack of statistical power.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

R1.0. Could the authors please check all references in the body of the manuscript to make sure they correspond with the correct reference in the reference list.

Response: We have checked each of the references and synchronized them to the reference list. Additionally, we have reduced the number of references from 55 to 40, prioritizing substantive and/or more recent citations.

R1.1. In the Abstract one of the odds ratios appears to report the lower bound of the 95% CI as the odds ratio (bisexual identity associated with RSOD (OR 1.02 95% CI 1.02). Please check this one.

Response: We have corrected the abstract and would like to thank the reviewer for spotting this.

R1.2. Article focus: regarding the second point about alcohol, please spell out what is mixed. While I get what the authors mean, it is unclear.

Response: This has been rewritten for clarity: ‘The extent to which gender modifies this association is not known however, which is a clear priority for future research’ (p.13).

R1.3. In the Methods section, I don't understand what is meant by "ethnic minority groups to N=1000 per group". Could the authors please explain this a bit more clearly if possible?

Response: We have rewritten to make this clear: ‘Socio-economically deprived schools were over-sampled by a factor 1.5 and ethnic minorities to achieve N=1000 per ethnic group.’ (p.6)

R1.4. In the Results, there is a % sign missing after 1.3 LG, 2.2% GB.

Response: This has been corrected, thank you.

R1.5. I would avoid using 'B' and 'L' and 'G' on their own and would opt instead to spell these words out when used on their own. LG and GB and LGB are fine however.

Response: We have adopted this suggestion.

R1.6. In the Results, there are several instances where the authors should spell out who they are comparing LG and bisexual respondents to so it is more clear. e.g. "Lesbian and gay participants

were more than twice as likely to have a history of cigarette smoking compared to heterosexual participants".

Response: This is an excellent suggestion which has been addressed in the article summary and results section.

R1.7. Were any comparisons made between LG and bisexual participants? If not, why not? Could the authors please cite as a limitation. I suspect it is because of a lack of statistical power.

Response: Comparisons were not made between these groups, for several reasons: our research question was focused on LGB or LG identity rather than LG vs. B identity, low statistical power, and simplifying presentation by limiting the number of comparisons made. This has now been acknowledged in the discussion section.